
Crowd Computation: Organizing Information during Mass Disruption Events

Kate Starbird

University of Colorado Boulder
ATLAS Institute
catharine.starbird@colorado.edu

Abstract

This research examines large-scale human interaction occurring through social media during times of mass disruption, seeking to understand how the connected crowd acts to organize a flood of data moving through those platforms into useful information resources. The work combines empirical analysis of social media communication, interviews, and participant observation to explore how people work to organize information and how they use social media platforms to organize themselves to do this work. Synthesizing findings from four distinct, yet interrelated studies, this research progresses towards a new conceptualization of the distributed, connected work of organizing information during mass disruption events.

Copyright is held by the author/owner(s).
CSCW'12, February 11–15, 2012, Seattle, Washington, USA.
ACM 978-1-4503-1051-2/12/02.

Keywords

Collective intelligence, crisis informatics, crowdsourcing, distributed cognition, social media

ACM Classification Keywords

H.5.3 Groups & Organization Interfaces—collaborative computing, computer-supported cooperative work, K.4.2 Social Issues

General Terms

Human factors, design

Introduction

Social media are playing host to new, digital forms [5] of the social convergence behavior long known to occur in the wake of mass emergencies and disasters [1,2]. This research intends to take an in-depth look at the ways in which social media users are processing a flood of data into useful information during mass disruption events. The inquiry extends from a perspective of *collective intelligence*, where users act collectively through their activity within information spaces to help organize data and create useful informational resources for responders and the public.

Background

Sociologists of disaster have repeatedly shown that after disaster events spontaneous volunteers converge on the scene and begin to offer help [1,2]. People are now converging via social media during mass disruption events, e.g. [5,6], participating in many of the same *informational convergence* activities documented in the pre-digital world [2]. Social media are being used for new types of information activities as well, e.g. to provide information to others during times of crisis [5], to help coordinate relief activities [7], and as places where digital convergers gather to integrate information [6].

The Problem of Informational Convergence

Informational convergence can be viewed as a problem. In the Disaster Relief 2.0 Report [3] humanitarian responders claimed that, during the aftermath of the Haiti earthquake, data generated on social media by those affected as well as digital volunteers that converged there added to responders' existing difficulties with processing large amounts of data.

There are four main components to this information-processing problem. First, a noisy information space must be filtered so people can find relevant information. Second, information via social media loses context and must be re-contextualized to restore its original meaning or give it new significance. Third, misinformation and disinformation can be dangerous during mass disruption events, so information must be verified. Finally, for information to be aggregated, searched, and integrated into other resources, it must be formatted or *structured* in a systematic way. Dealing with these problems requires a substantial amount of work, some of which is already taking place.

The Potential of Informational Convergence

One way that data are being filtered, re-contextualized, verified and structured is through the collective activity of the social media-connected crowd. During the early aftermath of the 2010 Haiti earthquake, several digital volunteer groups mobilized in attempts to help those affected and those responding [3]. For example, one community of remote volunteers came together to create and improve publicly available maps of affected areas, and another used an Ushahidi instance to collect, process, verify, and map citizen reports of damage and immediate needs.

Proposed Research

Informational convergence in the digital age presents new challenges and new opportunities for aggregating, processing, and utilizing information during mass disruption events. Though some are working on machine computational solutions to address these issues, these alone are not yet adequate to meet these new challenges. Disaster responders and the affected public need, and are currently developing and improvising, new strategies for filtering, re-contextualizing, verifying, and structuring this flood of data. This research asks:

How does the social media-connected crowd act to organize the flood of data moving through those platforms into useful information resources?

This broad question encompasses many different types of activities and possibilities for information processing in social media spaces. To approach a more detailed understanding of this diverse space, I ask three sub-questions:

RQ1. How can useful information be extracted from social media updates? How can people be taught to make their SM communications machine-readable?

RQ2. How do remote individuals use social media to organize themselves to process information during mass disruption events?

RQ3. How does the larger crowd act to organize information through individual actions within the social media space? How can we derive meaning from the collective "work" of the crowd?

By examining both the purposeful activity of digital volunteers and the ambient information-organizing activities of the connected crowd, I aim to understand how people are currently helping to organize information using social media tools and to inform solutions that better leverage this behavior.

Studies

The research consists of four separate studies. Though there is some temporal progression between the separate studies, they are not considered stages, but distinct and interacting investigations. Studies 1, 2 & 4 have been reported on in part in previously published or accepted work. Moving forward, I will expand and further synthesize previous findings, complete the studies, and tie findings together into a larger framework.

Study 1: Tweak the Tweet:

Introducing an Innovation for Structuring Information
Addressing RQ1, Study 1 involves an ongoing effort to examine the rationale, implementation, adoption, and use of the Tweak the Tweet (TtT) crisis-reporting syntax [8]. This is an innovation we proposed for the Twitter platform that allows users to make their Twitter

communications (tweets) machine-readable. An early portion of this study has been accepted for publication. Within my dissertation, I intend to explain how this initial project informed and evolved into the much larger research effort described here.

Study 2: Self-Organizing by Digital Volunteers

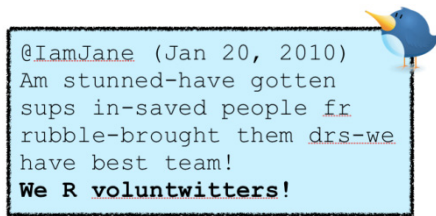
In our initial deploy of TtT, during the early aftermath of the catastrophic earthquake that struck Haiti on January 12, 2010, we were surprised to find that the syntax was adopted not by those on the ground during the event, but by remote individuals—digital volunteers—for whom syntax use was one tool and activity within a wide range of volunteer work. Addressing RQ2, Study 2 focuses on the activities of these digital volunteers ("voluntweeters"), exploring their motivations, resources and activities, describing how features of the technical environment offered structure for self-organizing, and considering how the Twitter platform enabled collective action [7].

Study 3: Humanity Road: An Emerging-Expanding Organization of Digital Volunteers

Study 3 extends the examination of the self-organizing of digital volunteerism to a nascent organization, *Humanity Road*, that emerged from coalescing digital volunteer activities on Twitter and elsewhere during the Haiti emergency and relief period. Like Study 2, Study 3 addresses RQ2, but investigates the digital volunteers at a later stage of organizational development.

Study 4: Using the Crowd as a Collaborative Filter

Study 4 is designed to address RQ3, approaching the information-processing problem from a perspective that considers the whole crowd as potential "workers." In this study, I explore the possibility of leveraging our understanding of information convergence behavior



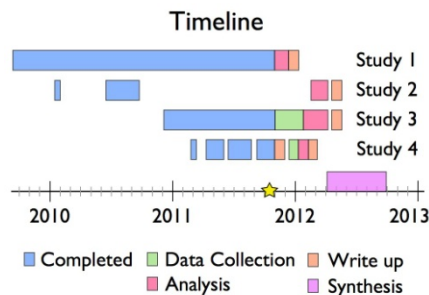
Tweak the Tweet Crisis Reporting Syntax

Prescriptive Tweet

@EPIC: To report a need: #eventtag
#need [list your needs, e.g: rescue,
or food & water] #loc [location or
GPS coords] #contact [email, @ or #]

Example Tweet

@barbaraslavin: #haiti #needs food
H20 #name Marie Girard #loc
Institut Franse, Ruelle Robin P-a-P
#info 5555-5555 #source @Delmon TtT



during crisis—specifically the patterns of recommending and routing information, following, and friending—to glean meaning from the aggregated micro-actions of social media users.

Analysis

These studies incorporate empirical analysis of Twitter data, interviews, and participant-observation from within a digital volunteer community. Study 4 also includes an effort to use machine learning techniques to interpret organization of the information space from the collective activity of the crowd. After demonstrating how existing frameworks of crowdsourcing, collective intelligence, and human computation fail to fully encompass and highlight the salient features of crowd work in the context of mass disruption, I plan to offer an alternative perspective through which to view these activities. I will explore the use of distributed cognition [4] to derive a conceptualization of crowd work as a connected, cognitive system. Synthesizing findings from the four studies, this work progresses towards a new framework, *crowd computation*, for the distributed, connected work of organizing information during mass disruption events.

CSCW Doctoral Colloquium

Crowdsourcing is emerging as a popular new research area. My research aims to describe and better understand some of the ways in which the connected crowd is doing work. By attending the CSCW Doctoral Colloquium, I hope to work through how to conceptualize these phenomena both within (or perhaps outside) the crowdsourcing framework, and through other, more fully developed CSCW perspectives. I also hope to gain insight into how to best communicate this understanding to other scholars in the CSCW field. Finally, I hope that

by attending the colloquium I will create connections with other future CSCW scholars who can enrich and support my intellectual development within this community.

Acknowledgements: This work is supported by a U.S. National Science Foundation Graduate Research Fellowship.

References

- [1] Dynes, RR. *Organized Behavior in Disaster*. Heath, 1970.
- [2] Fritz, C. E. & Mathewson, J. H. *Convergence Behavior in Disasters: A Problem in Social Control*, Committee on Disaster Studies, National Academy of Sciences, National Research Council, Wash. DC, 1957.
- [3] Harvard Humanitarian Initiative. *Disaster Relief 2.0: The Future of Information Sharing in Humanitarian Emergencies*. Washington, D.C. and Berkshire, UK: UN Foundation & Vodafone Foundation Technology Partnership, 2011.
- [4] Hutchins, E. *Cognition in the Wild*. MIT Press, 1995.
- [5] Palen, L & Liu, S. B. Citizen Communications in Crisis: Anticipating a Future of ICT-Supported Participation, *Proc. of CHI 2007*, 727-736.
- [6] Qu, Y., Wu, P. & Wang, X. Online Community Response to Major Disaster: A Case Study of Tianya Forum in the 2008 China Earthquake. In *Proc 42nd Hawaii Int'l Conf. on System Sciences, 2009*, 1-11.
- [7] Starbird, K., & Palen, L. 'Voluntweeters': Self-organizing by digital volunteers in times of crisis, *Proc of CHI 2011*.
- [8] Starbird, K., & Stamberger, J. Tweak the Tweet: Leveraging microblogging proliferation with a prescriptive syntax to support citizen reporting, *Proc. of ISCRAM 2010*.